A Scalable Inference Method For Large Dynamic Economic Systems

Anonymous Author(s)
Affiliation
Address
`email`

Abstract

The nature of available economic data has changed fundamentally in the last decade due to the economy's digitisation. With the prevalence of often black box data-driven machine learning methods, there is a necessity to develop interpretable machine learning methods that can conduct econometric inference, helping policymakers leverage the new nature of economic data. We therefore present a novel Variational Bayesian Inference approach to incorporate a time-varying parameter auto-regressive model which is scalable for big data. Our model is applied to a large blockchain dataset containing prices, transactions of individual actors, analyzing transactional flows and price movements on a very granular level. The model is extendable to any dataset which can be modelled as a dynamical system. We further improve the simple state-space modelling by introducing non-linearities in the forward model with the help of machine learning architectures.

1 Introduction

The digitisation of the economy and society had a profound impact on how researchers study economic interactions. The availability of enormous granular datasets has led to innovations in data-driven machine learning approaches, excelling in forecasting. However, the opaqueness of these approaches has been proven to limit the usefulness to policymakers. We thus develop a scalable inference model which combines innovations of popular machine learning methods such as variational approaches and non-linear forecasting methods with interpretable parametric models to study economic interactions at scale. We apply the model to study an emerging part of this new digital economy.

A prime example of this digitisation is the introduction of decentralised digital currencies which had a big impact on the global financial system. Blockchain, by using decentralised architecture, has been a disruptive innovation in terms of global transactions. With increasing interest in cryptocurrency trading, especially in ERC20 tokens, and increase in the number of exchanges providing a platform for the users to trade with ease, cryptocurrency trading has now become a complex economic system with complicated dynamics between price, trading volume and tokenflow.

Due to a large number of transactions and with every transaction being accessible online and traceable to some extent [1], we are left with granular data which can be used for extracting valuable new information about this ecosystem, like, interactions between on-chain (transaction amount and token inflow) and off-chain (trade data like closing price) ecosystems. Both these features being complementary, create a complex dynamical system similar to what we see in traditional financial markets. This granular dataset, in combination with our model, provides a data-driven approach to inform policymakers and can provide valuable information when creating new regulations to protect investors in this yet largely unregulated market.

Previous works have successfully modelled cryptocurrencies as a dynamical system using Time-Varying Parameter - Vector Autoregression (TVP-VAR) [2, 3]. They note that a time-varying regression coefficient could ideally model this ecosystem. As noted by Nadler et al. [3] and Haldane [4], the economic model needs to scale to deal with big data appropriately. We build on the

previous research and improve the Bayesian inference in a way that is scalable for big data. We introduce a novel algorithm, Time-Varying Parameter - Vector Autoregression - Variational Inference (TVP-VAR-VI) to tackle this problem. We test the proposed approach on synthetic data to validate and perform benchmarking. We then extended it to real-world blockchain data. We use this inference technique to interpret an otherwise unobservable parameter representing the dynamics of the blockchain ecosystem. The resulting vector parameter in latent space is used to visualise and interpret the influence of on-chain transactions on trading data, pricing and market dynamics. We use Bayesian inference as opposed to a neural network, as a simplistic NN forecasting architecture would yield a uninterpretable weight matrix in latent space. The parameter estimated using our technique can be interpreted visually, which has been further discussed in Section 5.

We also challenge the simplistic state-space model for the evolution of the latent parameter. We propose a new technique to improve the state-space modelling of the ecosystem, by incorporating non-linearity in our modelling by using data-driven approaches and develop a neural network architecture called TVP-VARNet.

In short, the Bayesian inference model helps us in obtaining the latent dynamic parameter which describes the interrelationship of economic variables dynamically over time whose structure economists can exploit for policy analysis. Our methodology is especially relevant for very large and granular datasets which would be computationally prohibitive for many established econometric models, since our model is a scalable algorithm for big data. Extending TVP-VAR-VI, we develop TVP-VARNet model which helps in overcoming the shortcoming of the simplistic state-space to forecast this latent parameter in case of out-of-sample n-step predictions.

2   Background

In the following, we will denote with $x_t \in \mathbb{R}^n$, the time evolving latent variables and $y_t \in \mathbb{R}^m$ as observations. $\mathcal{M}$ denotes the operator used to evolve the latent variables in time

$$x_t = \mathcal{M}(x_{t-1}) + \epsilon_t \tag{1}$$

where ϵ_t is Gaussian background error with covariance matrix Q. Finally, $\mathcal{H}: \mathbb{R}^n \to \mathbb{R}^m$ will denote the operator which maps the state variables to the observations.

$$y_t = \mathcal{H}(x_t) + \epsilon_o(t) \tag{2}$$

where $\epsilon_o(t)$ is Gaussian model error with with covariance matrix R.

2.1   Bayesian Inference

The objective of Bayesian inference is to update the belief of latent space (analysis, x_t^a) by ingesting the observation (y_t) and prior/model information (x_t^b) at each time step, t. The prior distribution here is the PDF of the latent state when no observations are available, x_t^b (i.e. output of the model (1)). The likelihood is the observations we get using the latent space in (2). Whereas, the posterior represents the final updated latent state, after considering observations [5]. Modelling these PDFs as gaussian process, with prior mean as x_t^b, covariance Q and likelihood mean as $\mathcal{H}(x_t)$, covariance R, we get

$$P(x_t|y_t, x_t^b) \propto \exp{-\frac{1}{2}\big[(x_t - x_t^b)^{\mathrm{T}} Q_t^{-1}(x_t - x_t^b) + (y_t - \mathcal{H}(x_t))^{\mathrm{T}} R_t^{-1}(y_t - \mathcal{H}(x_t))\big]} \tag{3}$$

We seek a solution maximising the posterior probability, $P(x_t|y_t, x_t^b)$ which means minimising the negative log likelihood, $-\ln P(x|y, x_b)$ or $J(x_t^a)$

$$
\begin{aligned}
J(x_t^a) &= (x_t - x_t^b)^{\mathrm{T}} Q_t^{-1}(x_t - x_t^b) + (y_t - \mathcal{H}(x_t))^{\mathrm{T}} R_t^{-1}(y_t - \mathcal{H}(x_t)) \\
&= J_b(x_t) + J_o(x_t)
\end{aligned}
\tag{4}
$$

Proven techniques for such inference schemes are the Kalman filter and Variational inference. The Kalman filter [6] approaches the problem statement sequentially by analytically solving the cost function and calculating a weighted average, known as Kalman Gain. In contrast, Variational inference method minimises the same cost function that reduces the analysis gap between both time distributed observations and model solution [7], thus working in a continuous way.

In the later sections, we will be reporting results of both Kalman and Variational inference methods, modified and implemented to suit TVP-VAR modelling. This is further extended to the blockchain data to understand the latent state dynamics.

2.2 Ethereum and ERC20 tokens

Ethereum is a specialised decentralised network that, along with recording transactions on blockchains, allows the creation of smart contracts. This gives an environment to create decentralised applications (DAPPs). A detailed description of Ethereum's architecture in available in Wood et al. [8].

ERC20 Tokens An Ethereum Token is a digital, blockchain-based asset created on top of the Ethereum network created using a smart contract. These tokens serve as proof of ownership of an amount of gold or a house. Figure 1 revealed that the market capital of ERC-20 tokens now represents
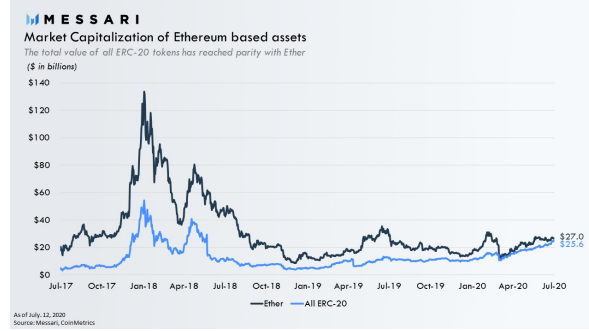

Figure 1: Market Capitalisation, Source: Messar [9]

49% of total assets on Ethereum. This increased share incentivizes economic activity using ERC20 tokens, providing excellent avenues for researching the trading activities and dynamics existing within this ecosystem.

Exchanges Cryptocurrency exchanges are online platforms where customers can trade one kind of digital asset or currency for another based on the market value of the given assets. These exchanges are intermediary between buyers and sellers of the cryptocurrency similar to the traditional stock exchange platforms. The most popular exchanges are currently Binance, GDAX, Poloniex etc [10]. Hence, we can see that these exchanges serve an integral part of the trading ecosystem in blockchain. These exchanges together process nearly $10 billion [11], playing a major role in driving this ecosystem.

3 Blockchain Ecosystem: Econophysical Analysis

3.1 Dynamic system formulation

We model the on-chain (transaction flows and amounts) and off-chain trading activities on blockchain exchanges as a Time Varying Parameter Vector- Autoregressive model (TVP-VAR). TVP-VAR embodies a system which has a set of vector autoregressive coefficients of time-varying nature. TVP-VAR has been used to model price dynamics of various cryptocurrencies in other literature [12] as well. The dynamic relation in the blockchain ecosystem can be time varying and thus TVP VAR helps us in accommodating this shift by varying coefficients, allowing the model parameters to vary across time. The multivariate lagged VAR is

$$y_t = \phi_1 y_{t-1} + ... + \phi_l y_{t-l} + \mu_t + \epsilon_t \sigma_t$$

where y_t is $q \times 1$ vector of observations, μ_t a vector of means and ϕ_l is a $q \times q$ coefficient matrix and l is the lag length. The model can be expressed in compact notation using $X_t = [y'_{t-1}, ..., y'_{t-l}, 1]'$ and $\phi = [\phi_1, ..., \phi_l, \mu]'$. With $K = (ql + 1)$, we get X_t of dimension $K \times 1$ and ϕ as $K \times q$. We further define $x_t = I \otimes X_t$ using Kronecker product $\otimes$ and I as $q \times q$ identity matrix and define $\beta_t = vec(\phi)$, $vec()$ stacking the columns of a matrix, dim $Kq \times 1$. This allows us to rewrite the TVP-VAR in compact notation. Further details can be found in [13, 14]:

$$y_t = x'_t \beta_t + \sigma \tag{5}$$
$$\beta_t = F\beta_{t-1} + v_t \tag{6}$$

Where v_t and σ are zero mean error co-variances with $v_t \sim N(0, Q_t)$ and $\sigma \sim N(0, R)$. β_t is the time-varying coefficient or the latent parameter that define the dynamic of the economic system. F is the forward model, similar to M in (1) to evolve β in time.

3.2 Link to Bayesian inference

A lot of times, the state representing interrelationship may evolve with time, however, this latent state can be unknown and not directly observable. These have to be inferred from the measurements/observations, which may be very sparse and noisy. Bayesian inference can help in analysing this latent space.

TVP-VAR with Bayesian inference is used in estimating the latent blockchain dynamics, β in (5). This method enables the use of both sparse and highly aggregated data effectively to infer dynamics to analyse the effect of trading actions on blockchain on-chain activities or how token flow effect the pricing actions on exchanges.

We can see that the (5) and (6) relate to the state-space (1) and (2) discussed above. Tweaking the state-space equations can give us a TVP-VAR formulation of our econophysical system and hence modified inference approaches can help us determine the unobserved state variable, i.e. gives us insight to the interrelationship between economic variables like x and y.

4 Our proposed TVP-VAR-VI: Variational Inference

Native variational data inference needs to be adapted for modelling the state-space economic time-varying model to incorporate observations and interpret the unobserved model parameter, β. We propose a new methodology to combine TVP-VAR with variational inference, formulating a novel algorithm called TVP-VAR-VI. Rewriting (4) compactly and adapting it to be inline with TVP-VAR equations (5),(6) we get a cost function that reduces the analysis gap between time distributed observations and model solution.

$$
J(\beta_t) = ||\beta_t^b - \beta_t||_{Q_t^{-1}} + \sum_{j=t}^{j=t+\Delta t} ||y_j - X_j\beta_j||_{\sigma^{-1}},
$$
$$
\beta_t^{opt} = \arg\min_{\beta_t} J(\beta_t)
$$
(7)

We minimise the cost function in (7) so as to incorporate observations, y and exogenous input X, to infer latent parameter β. Further details can be found in previous work by Nadler et al. [3].

In (7), following notations are used:

- Δt is window size, forecasts from this time window are used for optimising the initial value.
- 4D Variational approach is performed over a time window $(t, t + \Delta t)$, where β_t^{opt} is the optimised initial beta value for start of the window at t. Optimising (7) gives us β^{opt} every Δt time steps.
- F, model forecast is identity, to evolved β^{opt} as a random walk process. We later try to learn the F operator, i.e evolution of beta using deep learning.
- β^b is the background/prior beta, updated to β^{opt} as observations are ingested.

When $\Delta t = 0$, the cost function becomes that of 3D-Var (as time window, the 4th dimension in 4D no longer exists) and it minimises the function considering observations only that time. We propose the following algorithm combining TVP-VAR with variational approach.

Algorithm 1: TVP-VAR-VI

1: Initialisation β_0^b, Q_0, σ
2: **for** t = range(0, T, Δt) **do**   // t takes values like 0, Δt, 2Δt..
3:     $\beta_t = \arg\min J(\beta)$ // minimise 7 using a routine
4:     **for** i = 1..Δt **do**
5:         $\beta_{t+i} = F\beta_{t+i-1}$
6:         $Y_{t+i}^{hat} = x_{t+i}\beta_{t+i}$
7:     **end for**
8:     $\beta_t^b = \beta_{t+\Delta t}$
9: **end for**

As we see, there is no update step similar to Kalman (calculation of Kalman gain matrix); beta is inferred using iterative minimisation of the cost.

It is important to note that we need to update the background beta, β^b to optimal β^{opt} obtained from the minimisation and use this updated β^b as a starting point prior for the optimiser. Convergence is slower if a random starting point prior is used.

4.1 Implementation details

A challenging part of implementing TVP-VAR-VI was finding a way to computationally minimise the function in (7). We used automatic differentiation provided by TensorFlow. By calculating the function gradient from a graph, it allows us to free ourselves from deriving the gradient by hand, which is often infeasible and complex.

We empirically observed that Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) [15] optimiser worked best for us, LBFGS converged much faster (on an average 14% reduction in execution time) and over performed prevalent first order optimisers like Adam, SGD by reducing Mean squared error (MSE) in order of magnitude of 5.

L-BFGS, being one of the effective Quasi-Newton methods, gives us the ability to incorporate second order information (by estimating the Hessian) and hence could be the reason for optimal and better convergence given its ability to represent complex structures effectively.

5 Experiments

5.1 Synthetic data

We start with a well-behaved TVP-VAR synthetic data to evaluate the accuracy and scalability of the two approaches for big data as it gives us the ground truth for the latent parameters to validate our algorithms on. We then use the variational inference for off-chain and on-chain blockchain dataset after performing benchmarking. The data generation is similar to the one in [14],

$$y_t = X_t'\gamma_t + e_{1,t}$$
$$\gamma_t = \gamma_{t-1} + e_{2,t} + e_{3,t}$$

(8)

where γ represents the time varying state variable (β), and the errors, $e_{1..3}$ are assumed to be Gaussian white noise. Data matrix X_t is generated as standardised i.i.d process. Note that the forward model operator for γ is identity here. Figure 2 shows the sample beta and forecast plots estimated by VP-VAR-VI, i.e. our Variational approach.
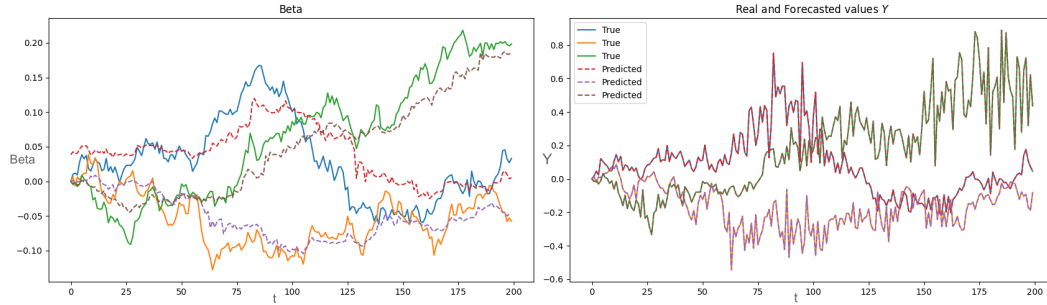

Figure 2: Window size 1, **TVP-VAR-VI**

We use the Mean square forecast error as our assessment metric. We implemented both the algorithms, TVP-VAR - Kalman and our novel TVP-VAR-VI using TensorFlow-gpu. This is done to ensure that the time difference noted was not due to difference in implementation details and technology used. Table 1 shows the results for both approaches. We can observe that as the problem size (n dimension ≥ 50) increases, the execution time increases for the Kalman by 10-40%. In contrast, for variational, the increase is hardly 1% for large problem size. With the increase in dimension, we see an increase in MSE for Variational, however the magnitude order (10^{-3}) remains the same, justifying the slight increase in MSE trade-off for lower execution time.

This can also be explained by looking at time complexities for the worst-case scenario. Time Complexity for matrix inversion (Kalman Gain) is about $O(n^3)$ (n is the size of observation - N dim) whereas for function minimisation using LBFGS is $O(mn)$ where m is the size of hessian history stored in memory, which is a constant hyper-param for LBFGS.

Another observation is that on comparing the experiments ran for both 4D-Var (window > 1, represented by the entry that has a missing value for Kalman, window = 2)) and 3D-Var (window = 1), 4D-Var has longer execution time, suggesting difficulty to reach the minima. We believe the

5

Time dim	N dim	Window	Exec Time(sec)		MSE	
			Kalman	Variational	Kalman	Variational
1000	10	1	0.4	25.7	2.60E-03	3.50E-03
2000	20	1	13.5	50.69	4.10E-03	1.00E-03
6000	20	1	55.78	150.56	3.90E-03	8.00E-02
6000	50	1	208.8	198.8	2.90E-03	8.90E-03
6000	50	2		204.56		7.30E-02
10000	50	1	312.8	225.8	3.50E-03	6.30E-02
10000	65	1	425.54	238.99	3.78E-03	3.50E-03
20000	65	1	850.18	479.85	3.82E-03	4.30E-03
10000	75	1	937.88	247.53	4.70E-03	5.60E-03
10000	80	1	1338.4	248.54	4.30E-03	6.06E-03
10000	85	1	1919.32	265.87	4.90E-03	9.02E-03

Table 1: Comparing Kalman vs Variational, with red gradient denoting the increasing execution time

reason could be that adding future observations (as part of the summation of values in a window) in the cost function could lead to a non-convex function with many local minima.

5.2 Blockchain data

After experimenting with well behaved TVP-VAR synthetic data, we empirically observed that as the dimension of data increases, the 3D-Var algorithm becomes more scalable; hence we use TVP-VAR-VI algorithm to interpolate latent parameters, β for the actual blockchain datasets.
The dataset of interest is the ERC20 tokens being traded on Ethereum. We deal with mainly two data sources, ERC20 tokens on-chain data; and off-chain data from exchanges.

On-Chain Data: We identify on-chain data as trades on Ethereum, amount of ERC20 tokens traded between blockchain addresses, timestamp of the transaction, number of transactions etc.
Off-Chain Data: We refer to the off-chain data as open-closing prices and trading volumes obtained using https://min-api.cryptocompare.com per exchange and ERC20 token.

Applying this methodology to economic data can help us infer the state of the system, how some features affect/drive other features. The main features used for blockchain dataset analysis were amount sent to/received by exchanges, number of transactions sent or received, the closing price of the token being analysed and volume from or to the token being traded.
For example, we tried to visually interpret the effect of on-chain token inflow on returns and vice-versa for BNB token on Binance exchange. As follows from the Figure 3, the dynamic is erratic
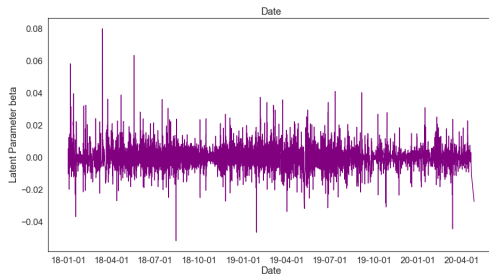

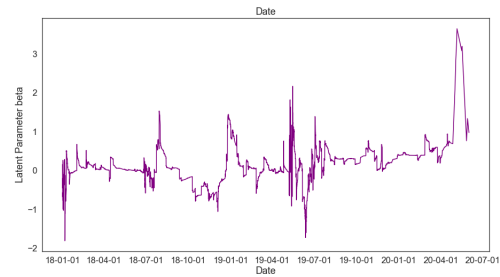Figure 3: Effect of amount to tokenflow on return for BNB token

Figure 4: Effect of return on token inflow for BNB Token

and chaotic for on-chain effect on trading prices, leading to the suspicion that the interaction is not structural. Prominent theories exist for traditional exchanges describing that inflow drives the price of the asset. It is very likely not the case for Ethereum ecosystem as suggested by our interpretation of the dynamic. However, we see a more prominent indication of the dynamic that represents the effect of returns on on-chain outflow/inflow of tokens when looking at Fig 4. This implies a one-sided relationship, with token inflows having an erratic and weak relation on returns; whereas returns having a structural impact on the inflow. These results were also corroborated by experiments for other ERC20 tokens like OMG and TRX.

6

This is possible evidence that the price and trading action probably attracts attention and drives the inflow of tokens towards a particular exchange or it could be due to arbitrageurs who benefit from price difference and cause the fluctuation of tokenflow [16]. However, on-chain flows don't drive prices. It could be that over time the impact of on-chain activity has become more decoupled from price movements on exchanges. This can serve as an important indication factor to policymakers, which can help them to create and shape future regulations concerning ERC20 tokens. It shows, for example, price manipulation is unlikely to be driven by activities on the chain; thus, regulation can focus on exchange specific risk factors such as orderbook fraud or sentimental actions.

6 Improving state-space modelling

The state-space model stated in (6) assumes that approximation of the beta evolution function, F is known. For the complex dynamics inherent in these economic systems, it is hard to know F, and with increasing dimensions, the non-linearities would become stronger; thus linear assumption of the forward model failing [17].

The TVP-VAR-VI would be able to update the prior estimate of the latent parameter when observations are available; however, the state-space model would fail for out-of-sample predictions.

As seen in plots above, there have been fluctuations in values of β that might not be driven only by the previous value and innovation but also a non-linear combination of some lagged values of β.

Running TVP-VAR-VI provides us with a complete state of time series of the latent parameter, β, with which we can obtain a surrogate forward model using machine learning tools.

Learning evolution of the unobserved β parameter has been challenging. We were trying to predict with a time series which had high volatility and was chaotic in nature. Also, multivariate-time series forecasting has been challenging ongoing research as it needs to appropriately leverage the dynamics of multiple variables [18]. Hence, different architectures were tried before settling on the final model.

6.1 TVP-VARNet Model

Figure 5 presents an overview of our proposed architecture, TVP-VARNet to model latent parameter dynamics of blockchain data. The architecture benefits from an ensemble of two neural networks,
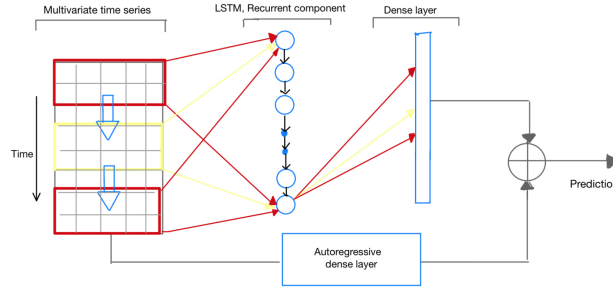

Figure 5: TVP-VARNet

an LSTM and an Autoregressive dense network with their outputs combined. The LSTMs has been successful due to its ability to capture and effectively exploit temporal dependencies, and take history into account for future state prediction.

However, as suspected by Lai et al. [18], LSTM could lead to an issue where the scale of outputs might not be sensitive to the scale of inputs. In the evolution of our unobserved β parameter, we see constant changes which are also in a non-periodic manner, resulting in poor forecasting of the model with only LSTM. To address this issue, similar in spirit to the LSTNet [18], we combine the final prediction of LSTM with a linear part, hence the model benefits from non-linear part having recurrent patterns through LSTM and a linear part through this component. By plotting the predictions on training and test set using only LSTM (see Figure 6) and using the TVP-VARNet (see Figure 7), we can understand the cruciality of the Autoregressive component. TVP-VARNet does a good job in generalising, by working appropriately for both kinds of evolution, chaotic and structured. Our architecture was able to adapt and forecast a satisfactory result for two-step ahead predictions.

It is important to note that the combination of TVP-VAR-VI and TVP-VARNet is crucial. TVP-VAR-VI provides us with a set of latent parameters to train our model, whereas TVP-VARNet helps us with the out-of-sample forecast. As and when new observations are available, TVP-VAR-VI can be
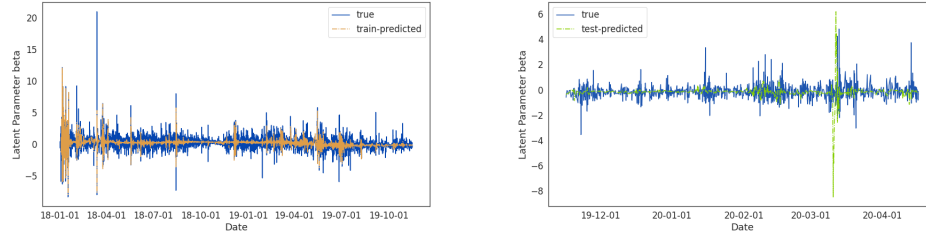
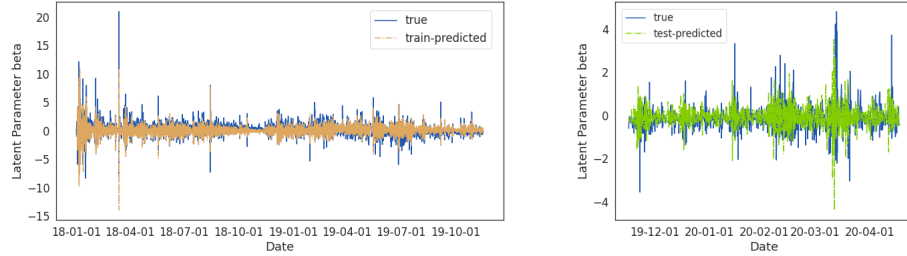Figure 6: LSTM model predictions for training and test set

Figure 7: TVP-VARNet model one-step predictions for training and test set

run to obtain updated β values to fine tune the TVP-VARNet hence giving us the ability to bring continuous correction to our machine learning model.

7 Conclusion, future work and social impact

The findings of this study could be understood as further validation of treating the cryptocurrency system as an econophysical system. By ingesting data from various on-chain transactions and trading information, we perform Bayesian inference to infer the latent time-series, which can significantly help in analysing the dynamics that exist in the economic system. We built a scalable novel TVP-VAR-VI algorithm. Before TVP-VAR-VI, the Kalman Filter approach was computationally expensive. We were able to reduce the execution time in the range of 10-40% as dimensions increased.

Our interpretation and results of the unobserved state are broadly consistent with conclusions drawn by Nadler et.al. The most prominent finding to emerge was the direction of interaction between on-chain token transaction/flows, and off-chain trading action was one-sided. Token inflow's effect on trading actions resulted in weak and chaotic interaction; however, has a high-persistent effect when reversing the roles. Trading price action has a considerate impact on driving the inflow of tokens for an exchange. We also observed that change in trading volume also had a structural effect on the returns.

One of the other significant contributions is challenging the latent state-space evolution and replacing it with a more nuanced, non-linear neural network model. The TVP-VARNet architecture models both kinds of latent dynamics: chaotic and structural. The relevance of having a surrogate forward model for the latent parameter evolution is supported by the fact that the out-of-sample forecast faces many hurdles with the simplistic state-space model. Our TVP-VARNet brings the possibility to forecast latent parameters for certain future time-steps effectively.

One of the directions for future work includes further analysis of this dynamic on highly granular order-book data. Ideally, these experiments can be replicated to any system that can be modelled as a TVP-VAR model, to understand their latent dynamics in a scalable fashion.

Social Impact The insight to the time-varying dynamical interrelationship between tokenflow, price movements and volume activities can give much-needed information about price and market dynamics in the blockchain ecosystem, which has been unexplored and unregulated. We believe that the traditional financial market rules do not apply to a peer-to-peer based financial market, hence can pave the way for blockchain analyst and policymakers to understand what factors might or might not affect said movements. This intern can help in avoiding fraudulent activities or arbitrage. Our work can we extended to any econophysical system modelled using TVP-VAR equations.

References

[1] Rainer Bohme, Nicolas Christin, Benjamin Edelman, and Tyler Moore. Bitcoin: Economics, technology, and governance. *Journal of Economic Perspectives*, 29(2):213–38, May 2015. doi: 10.1257/jep.29.2.213. URL `https://www.aeaweb.org/articles?id=10.1257/jep.29.2.213`.

[2] Christian Hotz-Behofsits, Florian Huber, and Thomas Otto Zorner. Predicting crypto-currencies using sparse non-gaussian state space models. *Journal of Forecasting*, 37(6):627–640, 2018.

[3] Philip Nadler, Rossella Arcucci, and Yi-Ke Guo. Data assimilation for parameter estimation in economic modelling. In *2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, pages 649–656. IEEE, 2019.

[4] Andrew G Haldane. Will big data keep its promise. *Speech at the Bank of England Data Analytics for Finance and Macro Research Centre, King's Business School*, 2018.

[5] Ross Bannister. Variational data assimilation background and methods. URL `http://www.met.reading.ac.uk/~darc/training/ecmwf_collaborative_training/Var-theory_vn02.pdf`.

[6] Rudolf E Kalman. On the general theory of control systems. In *Proceedings First International Conference on Automatic Control, Moscow, USSR*, pages 481–492, 1960.

[7] Mark Asch, Marc Bocquet, and Maëlle Nodet. *Data Assimilation: Methods, Algorithms, and Applications*. Society for Industrial and Applied Mathematics. ISBN 978-1-61197-453-9 978-1-61197-454-6. doi: 10.1137/1.9781611974546. URL `http://epubs.siam.org/doi/book/10.1137/1.9781611974546`.

[8] Gavin Wood et al. Ethereum: A secure decentralised generalised transaction ledger.

[9] Crypto news, pricing, and research. URL `https://messari.io/`.

[10] Jake Frankenfield. Bitcoin exchange definition, Aug 2020. URL `https://www.investopedia.com/terms/b/bitcoin-exchange.asp`.

[11] Author Blockgeeks and Blockgeeks. What is the 0x project? the most comprehensive guide ever written, Apr 2020. URL `https://blockgeeks.com/guides/0x-project/`.

[12] Christian Hotz-Behofsits, Florian Huber, and Thomas Otto Zörner. Predicting crypto-currencies using sparse non-gaussian state space models. *Journal of Forecasting*, 37(6):627–640, 2018. doi: 10.1002/for.2524. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/for.2524`.

[13] Jouchi Nakajima et al. Time-varying parameter var model with stochastic volatility: An overview of methodology and empirical applications. Technical report, Institute for Monetary and Economic Studies, Bank of Japan, 2011.

[14] Philip Nadler, Rossella Arcucci, and Yi-Ke Guo. A Scalable Approach to Econometric Inference. page 10, .

[15] Jorge Nocedal and Stephen Wright. Numerical optimization second edition, 2006. URL `https://pages.mtu.edu/~struther/Courses/OLD/Sp2013/5630/Jorge_Nocedal_Numerical_optimization_267490.pdf`.

[16] Philip Nadler, Rossella Arcucci, and Yike Guo. An Econophysical Analysis of the Blockchain Ecosystem. page 16, .

[17] Suraj Pawar, Shady E Ahmed, Omer San, Adil Rasheed, and Ionel M Navon. Long short-term memory embedded nudging schemes for nonlinear data assimilation of geophysical flows. *arXiv preprint arXiv:2005.11296*, 2020.

[18] Guokun Lai, Wei-Cheng Chang, Yiming Yang, and Hanxiao Liu. Modeling long-and short-term temporal patterns with deep neural networks. In *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, pages 95–104, 2018.